2-26-2019

Comparison of Different Approaches to Evaluate and Explain Interviewer Effects

Geert Loosveldt
KU Leuven, geert.loosveldt@kuleuven.be

Celine Wuyts
KU Leuven, Celine.Wuyts@kuleuven.be

KU LEUVEN

A comparison of different approaches to evaluate and explain interviewer effects

Geert Loosveldt
Celine Wuyts
Caroline Vandenplas

Centre for Sociological Research

Outline

- Introduction
  - Conceptual issues
- Basic model
- Data
  - European Social Survey round 8
- Three procedures to evaluate interviewer effects
  - Two step procedure
  - Conditional random interviewer effect model
  - The mixed effect location scale model
  - Focus: impact of respondent characteristics

Introduction

- Interviewer effects:
  - differences between interviewers in their systematic effects on the respondent's answers
  - Additional variance

- Evaluation of interviewer effects:
  - Essential part of data quality assessment
  - Variance analysis
  - Intra class correlation: proportion of explained variance
  - Only one type of interviewer effect

The basic model

- Two level hierarchical data structure:
  - Respondents are nested within interviewers
  - Two level random intercept (null) model

$$Y_{ij} = \beta_{0j} + \varepsilon_{ij}$$

$$\beta_{0j} = \gamma_{00} + \mu_{0j}$$

$$Y_{ij} = \gamma_{00} + \mu_{0j} + \varepsilon_{ij}$$

- β_{0j} intercept for interviewer j
- ε_{ij} residual error term for respondent i; variance σ_e^2
- μ_{0j} an interviewer-specific part of the intercept ;variance σ_u^2

  - Interviewer effects= intra class correlation coefficient:

$$\rho_{\text{int}} = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2}$$

KU LEUVEN

The basic model (Cn't)

- Extension at the respondent level: respondent characteristics e.g. X_1

$$Y_{ij} = \gamma_{00} + \mu_{0j} + \beta_{1j} X_{1ij} + \varepsilon_{ij}$$

  - Explanation of the variability in the substantive dependent variable
  - Evaluation of interviewer effects:
    - Partial control for the differences between interviewers in the composition of the respondent group
  - No direct assessment of the impact of respondent characteristics on interviewer effects

The basic model (Cn't)

- Extension at the interviewer level: interviewer characteristics I_1 (e.g. experience, workload, …)

$$\beta_{0j} = \gamma_{00} + \gamma_{01} I_{1j} + \mu_{0j}$$

  - Explanation of the differences between interviewers (interviewer effects) concerning the random intercept
  - Assessment of interviewer characteristics on interviewer effects
- Integration:

$$Y_{ij} = \gamma_{00} + \mu_{0j} + \beta_{1j} X_{1ij} + \gamma_{01} I_{1j} + \varepsilon_{ij}$$

  - explanation of variance of a substantive variable
  - No direct assessment of the impact of respondent characteristics on interviewer effects

Data

- European Social Survey (Round 8)
  - 21 countries
  - 9 questions from two climate change and energy module and 6 questions from welfare attitude module (11-point scale)
  - Control variables: gender, age, language of interview is the respondent's home language (0= no; 1=yes), self reported degree of urbanization (1= Big City, …. , 5 =countryside)
  - Highest level of education (EISCED variable):
    - level 1: Lower secondary school;
    - level 2: Upper secondary education or advanced vocational education,
    - level 3: Tertiary education.

Data

- Preliminary analysis of intra interviewer correlations
  - 15 variables in 21 countries
  - Model: basic random intercept model

$$y_{ij} = b_0 + a_2(Education = 2) + a_3(Education = 3) + b_1 Age + b_2 Male$$

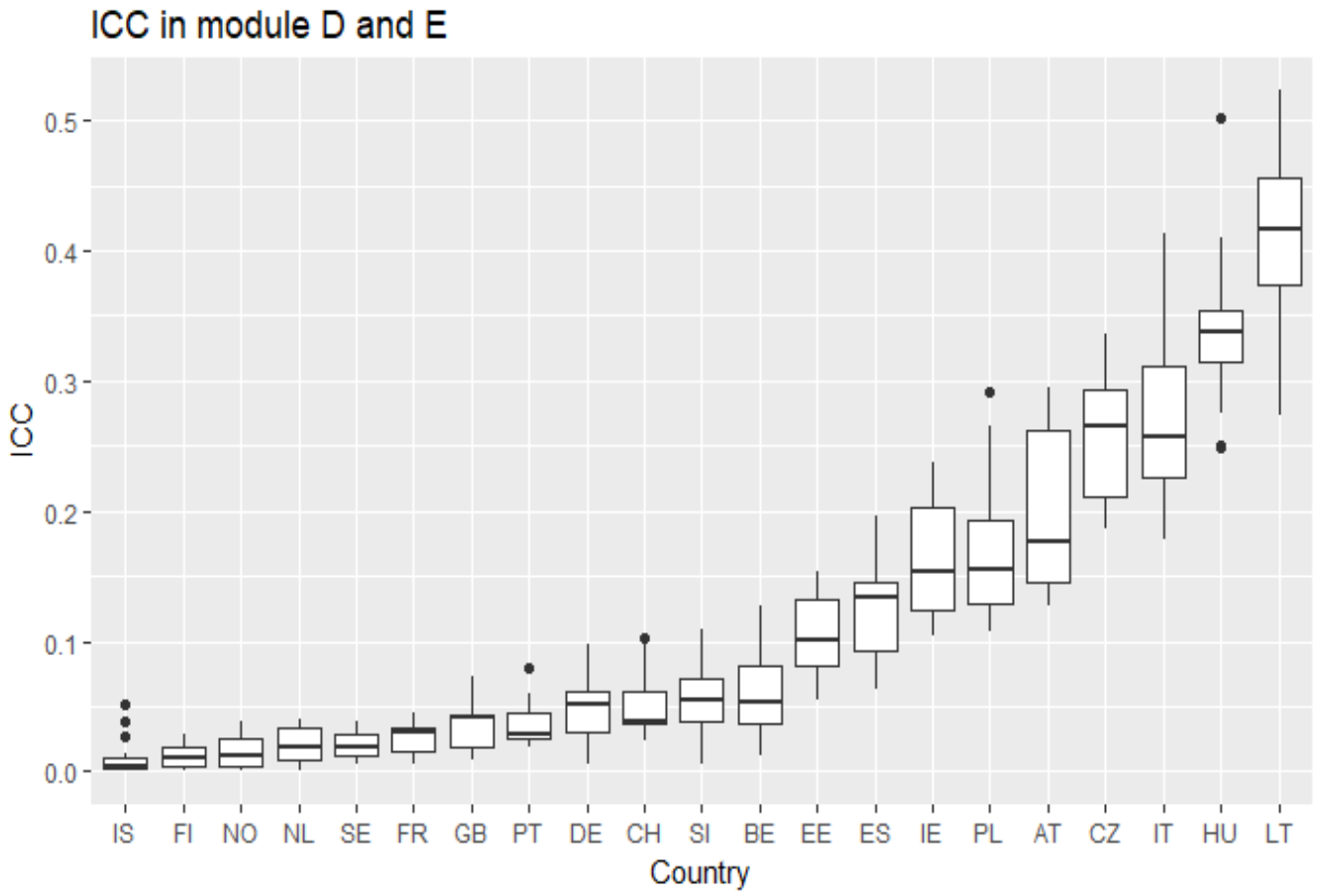
$$+ b_3 Same\ language + b_4 Domicile\ type + u_j + \varepsilon_{ij}$$

$$u_j \sim N(0, \sigma_u^2) \qquad \varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2)$$

  - Intra Interviewer correlations (IICs) for each variable in each country

Data

- Figure: Boxplots of the Intra Interviewer correlations for 15 questions per country.



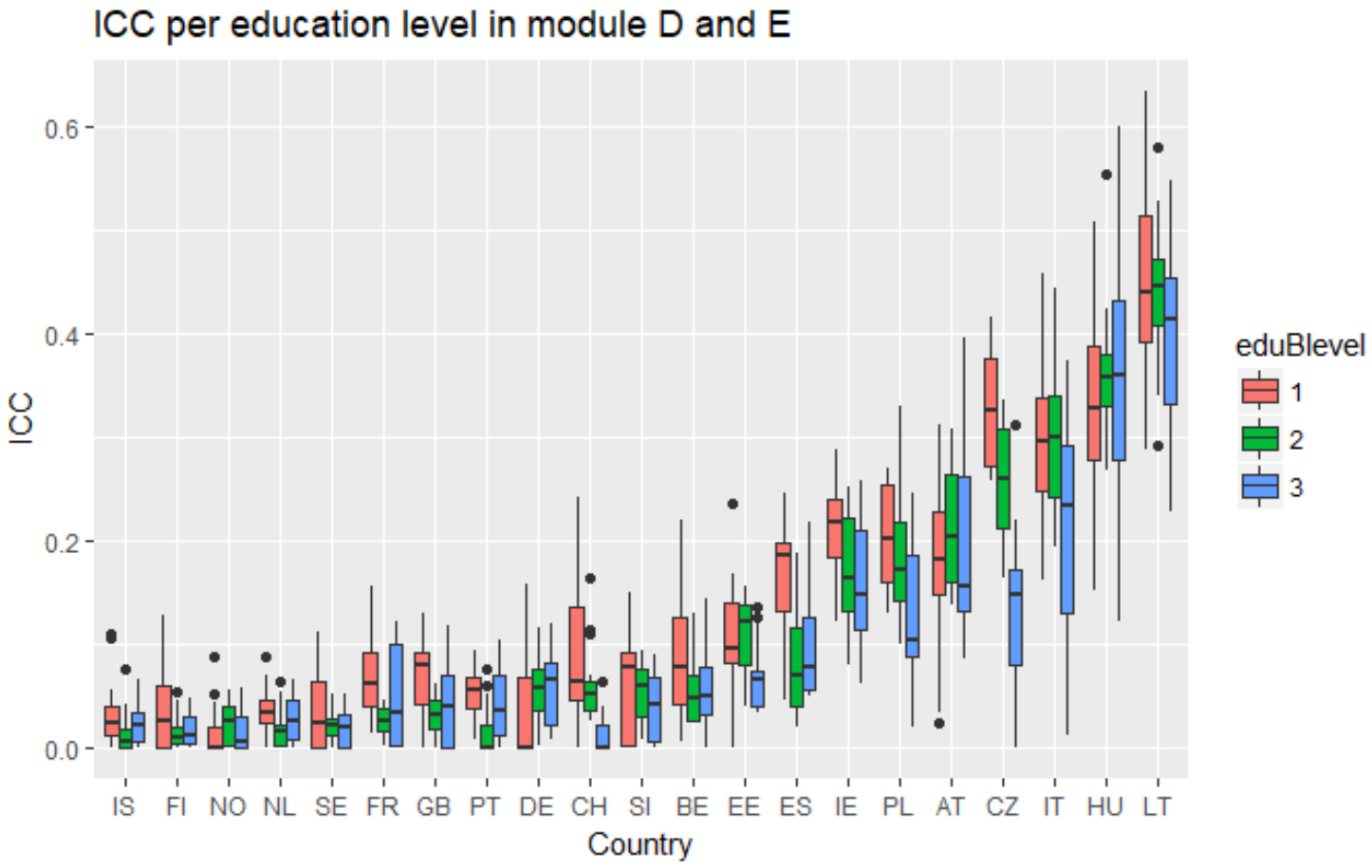
KU LEUVEN

Three procedures to evaluate interviewer effects and the impact of a respondent characteristic

- Respondent's educational level
- Expectation: more complex interactions→ higher interviewer effects

- Procedure 1: A two-step procedure
  - First step: Calculation of IIC's for 15 questions within 3 categories of educational level in each country
    - Basic model with control variables
    - Interviewers with at least three respondents in a given respondent group
    - Not the same interviewers in each respondent group
    - Number of IIC's: 15 questions x 3 levels x 21 countries= 945
    - Dataset with IICs as units:
      - With information about the question, educational level and country

- Figure: Boxplots of the IICs per education group and country
  - 12 countries: mean IIC in the lower group is 1,5 times the mean in the higher group



- Step 2: modelling of the IICs
  - Dependent variable: IICs
  - Independent variable: respondent characteristic(s) (R)
  - IICs are nested within questions and countries
  - Cross classified model (Question x Country)
    - Random intercept for country and question
  - Model:

$$IIC(question, country, edu) = b_0 + b_1 edu1 + b_2 edu2 + u_{0,country} + v_{0,question} + \varepsilon_{ij}$$

- Table: Fixed and random effects of the cross classified model with education for ICCS of the first procedure

```
Fixed effects:
              Estimate Std. Error
(Intercept)    13.7808     2.5582  ***
factor(eduB)1  -1.8281     0.4479  ***
factor(eduB)2  -3.5304     0.4479  ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1


Random effects:
 Groups   Name        Variance Std.Dev.
 CNTRY    (Intercept) 132.615  11.516
 TARGET   (Intercept)   1.934   1.391
 Residual              31.592   5.621
Number of obs: 945, groups:  CNTRY, 21; TARGET, 15
```

  - High mean IIC in the lower educated group (14%)
  - Significant decrease when education level increases
  - The variability of the intercept at the country level is larger than the variability between questions

- Procedure 2: Conditional random interviewer effect model
  - Conditionality is part of the initial model
  - For each variable
    - Random intercepts with variances within the categories of the respondent characteristic
    - The assumption of homogenous residual variance is relaxed
      - Residual variances within groups
    - Conditional IICs within each category
  - Test of differences between conditional IICs
    - Assessment of the relationship between respondent characteristic and the IICs

- Procedure 2: Model

$$y_{ij} = b_0 + a_2(Education = 2) + a_3(Education = 3) + b_1 Age + b_2 Male$$
$$+ b_3 Same\ language + b_4 Domicile\ type + u_{1j}(Education = 1)$$
$$+ u_{2j}(Education = 2) + u_{3j}(Education = 3) + \varepsilon_{ij}$$

with $\varepsilon_{ij} \sim N(0, \sigma^2_{\varepsilon 1})$ if $Education = 1$

$\varepsilon_{ij} \sim N(0, \sigma^2_{\varepsilon 2})$ if $Education = 2$

$\varepsilon_{ij} \sim N(0, \sigma^2_{\varepsilon 3})$ if $Education = 3$

$u_{1j} \sim N(0, \sigma^2_{u1}), u_{2j} \sim N(0, \sigma^2_{u2}), u_{3j} \sim N(0, \sigma^2_{u3})$

$$IIC(edu1) = \frac{\sigma^2_{u_1}}{\sigma^2_{u_1} + {\sigma_1}^2}$$

$$IIC(edu2) = \frac{\sigma^2_{u_2}}{\sigma^2_{u_2} + {\sigma_2}^2}$$

$$IIC(edu3) = \frac{\sigma^2_{u_3}}{\sigma^2_{u_{3j}} + {\sigma_3}^2}$$

- Conditional IICs: conditional variances of the intercepts and conditional residual variances

- Same information as in the two step procedure
  - Similar results for the descriptive analysis
  - Similar results for the modelling of the conditional IICs

- Homogeneity test of the covariance matrix
  - Each variable in each country (315 tests)
  - Null hypothesis: variance of the random intercept and the residual variance are equal across the three educational groups
  - Rejection of the null hypothesis in 47% of the test.
    - Variance components used to calculate the IICs are significantly (0,05) different for the three education levels

- Procedure 3: the mixed effect location scale model
  - Mean function (location part):
    - Random intercept: differences between interviewers concerning the mean

$$y_{ij} = b_0 + a_2(Education = 2) + a_3(Education = 3) + b_1 Age + b_2 Male$$

$$+ b_3 Same\ language + b_4 Domicile\ type + u_j + \varepsilon_{ij}$$

with $\varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2)$

$u_j \sim N(0, \sigma_u^2)$

  - Variance function (scale part):
    - Random residual variance: differences between interviewers concerning the residual variance

$$\ln(\sigma_\varepsilon^2) = \delta_{00} + \vartheta_{0j}$$

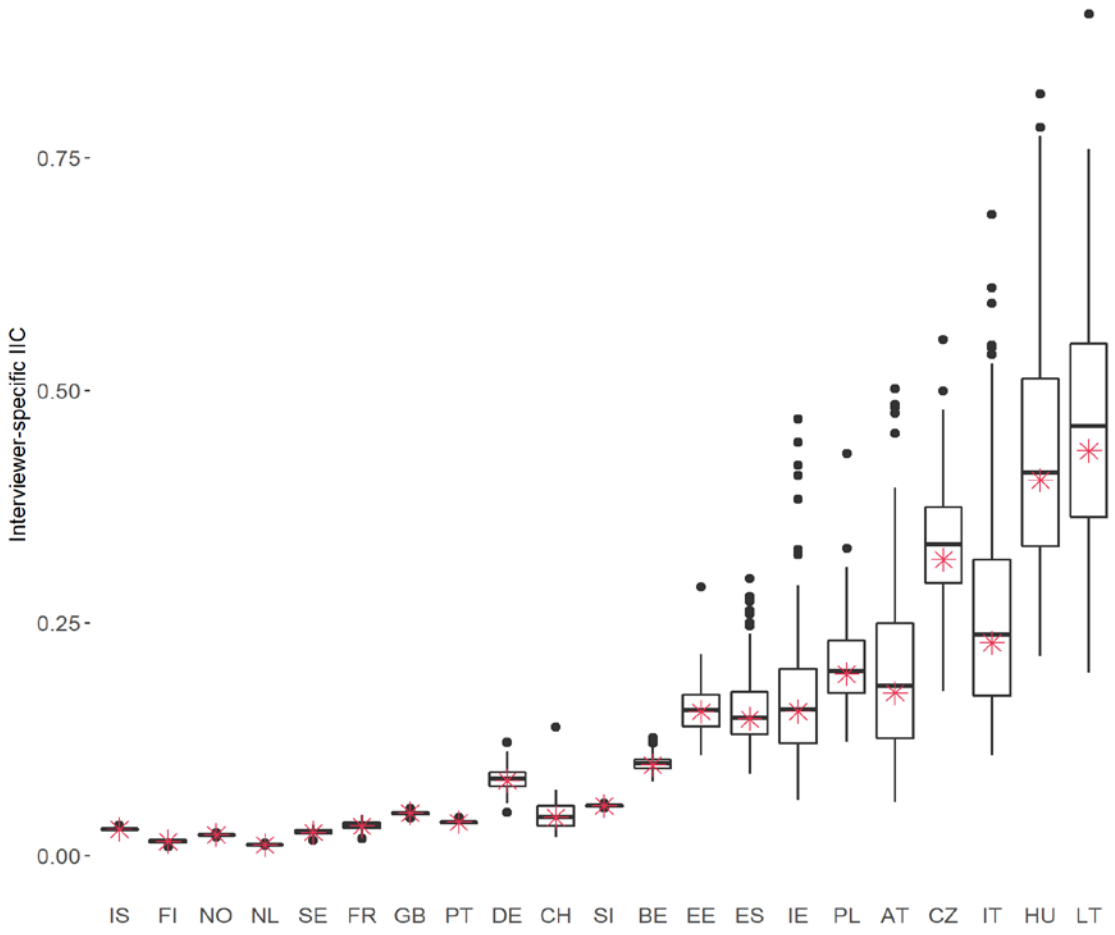
with $\vartheta_{0j} \sim N(0, \sigma_\vartheta^2)$ and $Cov(u_j, \vartheta_{0j}) = 0$

- Interviewer specific IICs:

$$IIC.j = \frac{\sigma_{\mu}^{2}}{\sigma_{\mu}^{2} + \exp(\delta_{00} + \vartheta_{0j})}$$

- SAS Proc NLMIXED
  - Starting values
- Illustration: 1 item
  - ‘personal responsibility to reduce climate change?’
  - Location: significant differences between interviewers for the intercept in 16 countries
  - Scale: significant differences between interviewers for the residual variance in 14 countries
  - Interviewers tend to vary in their mean response scores and in the variability of the responses

- Boxplots for the interviewer specific IICs for the item ‘personal responsibility to reduce climate change’



KU LEUVEN

- Link with respondent's educational level
  - Simple linear regression model at the interviewer level

$$\widehat{ICC}_j = \beta_0 + \beta_1 Proportion\ lower\ educated_j + \epsilon_j$$

with $\epsilon_{ij} \sim N(0, \sigma_\epsilon^2)$

  - Dependent variable: interviewer specific IIC for the item
  - Independent variable: proportion of lower educated respondents
  - In two countries: a significant negative slope
    - Higher proportion of lower educated respondents results in a smaller interviewer specific IIC.

Conclusion & discussion

- High interviewer effects in some countries
- Two step procedure and the conditional random intercept model
  - Conditional IICs within a limited number of respondent categories
  - Same results
  - Interviewer effects are higher for lower educated respondents
- The mixed effect location scale model
  - Computational demanding
  - Identification of interviewers with high ICCs
    - Response patterns?
  - Integration of respondents characteristics
    - Suggestions ?

KU LEUVEN

Thank You

KU LEUVEN